



# Inverse distributed hydrological modelling of Alpine catchments

H. Kunstmann, J. Krause, S. Mayr

## ► To cite this version:

H. Kunstmann, J. Krause, S. Mayr. Inverse distributed hydrological modelling of Alpine catchments. Hydrology and Earth System Sciences Discussions, 2006, 10 (3), pp.395-412. hal-00304868

**HAL Id: hal-00304868**

**<https://hal.science/hal-00304868>**

Submitted on 7 Jun 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Inverse distributed hydrological modelling of Alpine catchments

H. Kunstmann<sup>1</sup>, J. Krause<sup>2</sup>, and S. Mayr<sup>1</sup>

<sup>1</sup>Institute for Meteorology and Climate Research (IMK-IFU), Forschungszentrum Karlsruhe, Kreuzeckbahnstraße 19, 82467 Garmisch-Partenkirchen, Germany

<sup>2</sup>Institute of Geographic Sciences, Freie Universität Berlin, Malteserstraße 74–100, 12249 Berlin, Germany

Received: 6 October 2005 – Published in Hydrol. Earth Syst. Sci. Discuss.: 1 December 2005

Revised: 20 March 2006 – Accepted: 17 May 2006 – Published: 7 June 2006

**Abstract.** Even in physically based distributed hydrological models, various remaining parameters must be estimated for each sub-catchment. This can involve tremendous effort, especially when the number of sub-catchments is large and the applied hydrological model is computationally expensive. Automatic parameter estimation tools can significantly facilitate the calibration process. Hence, we combined the nonlinear parameter estimation tool PEST with the distributed hydrological model WaSiM. PEST is based on the Gauss-Marquardt-Levenberg method, a gradient-based nonlinear parameter estimation algorithm. WaSiM is a fully distributed hydrological model using physically based algorithms for most of the process descriptions.

WaSiM was applied to the alpine/prealpine Ammer River catchment (southern Germany, 710 km<sup>2</sup>) in a 100×100 m<sup>2</sup> horizontal resolution. The catchment is heterogeneous in terms of geology, pedology and land use and shows a complex orography (the difference of elevation is around 1600 m). Using the developed PEST-WaSiM interface, the hydrological model was calibrated by comparing simulated and observed runoff at eight gauges for the hydrologic year 1997 and validated for the hydrologic year 1993. For each sub-catchment four parameters had to be calibrated: the recession constants of direct runoff and interflow, the drainage density, and the hydraulic conductivity of the uppermost aquifer. Additionally, five snowmelt specific parameters were adjusted for the entire catchment. Altogether, 37 parameters had to be calibrated. Additional a priori information (e.g. from flood hydrograph analysis) narrowed the parameter space of the solutions and improved the non-uniqueness of the fitted values. A reasonable quality of fit was achieved. Discrepancies between modelled and observed runoff were also due to the small number of meteorological stations and corresponding interpolation artefacts in the orographically complex terrain. Application of a 2-dimensional numerical

groundwater model partly yielded a slight decrease of overall model performance when compared to a simple conceptual groundwater approach. Increased model complexity therefore did not yield in general increased model performance.

A detailed covariance analysis was performed allowing to derive confidence bounds for all estimated parameters. The correlation between the estimated parameters was in most cases negligible, showing that parameters were estimated independently from each other.

## 1 Motivation

Efficient parameter estimation techniques are a basic prerequisite for the successful application of distributed hydrological models to questions of sustainable water management. Parameter estimation must be performed before final assessments on water availability, flood risk or hydrological impact analysis of regional climate change can be addressed.

Even in physically based hydrological models, a set of parameters generally remains that must be calibrated. Lumped parameters (such as recession constants) must even be calibrated for each sub-catchment, which may involve considerable effort. Distributed and physically based hydrological models are usually much more CPU-time demanding than pure lumped models. This is particularly true for hydrological models that solve nonlinear equations (such as the Richards equation for infiltration in the unsaturated zone) and that couple to 2- or 3-dimensional groundwater models (like the hydrological model applied in this work, WaSiM). Here, systems of equations must be solved numerically and the effort increases with the number of horizontal grid points and the vertical resolution. Often, calibration and parameter estimation are performed by the usual “trial and error” method. Parameters are adjusted manually until simulated and observed river runoffs correspond well. Parameter estimation algorithms can facilitate this effort.

Correspondence to: H. Kunstmann  
(harald.kunstmann@imk.fzk.de)

For an optimization and automatization of the parameter estimation process, different algorithms have been investigated in hydrological modelling. Among the prominent ones that are looking for global optimal parameter sets, there is the Simulated Annealing Method (e.g. Press et al., 1992; Aarts and Korst, 1989; Abdulla et al., 1999), the Shuffled Complex Evolution Global Optimization Algorithm (SCE-UA) (Duan et al., 1992, 1994), the shuffled Complex Evolution Metropolis method (SCEM-UA and MOSCEM) (Vrugt et al., 2003a, b), and the multiobjective complex evolution algorithm (MOCOM) (Yapo et al., 1997). All of the above mentioned automatic parameter estimation methods, however, require a large number of model calls, usually in the order of several thousand. This limits their application to fast conceptual hydrological models or 1-dimensional Soil-Vegetation-Atmosphere-Transfer (SVAT) schemes, e.g. for the BATS scheme (Dickinson et al., 1993) by the Multiobjective Generalized Sensitivity Analysis (MOGSA) methodology by Bastidas et al. (1999).

When large, nonlinear and CPU time demanding models have to be calibrated, the number of model runs must be as small as possible. The Gauss-Marquardt-Levenberg method has the advantage that it can generally estimate parameters using fewer model runs than any other estimation method for nonlinear models (Doherty and Johnston, 2003; Doherty, 2004). It has the disadvantage that it is only a local search automatic parameter estimation tool, whose results may depend on the initial value of the search. The success of local search algorithms strongly depends on narrowing the solution space by additional a-priori information. Due to the comparatively small number of subcatchments in this study, parameter estimation approaches using regionalization concepts (as e.g. presented in Hundecha and Bardossy, 2004) could not be applied. We adapted the model-independent Parameter Estimation Tool PEST (Doherty, 2002) for automatic parameter estimation in this study. It is a public domain code that applies the Gauss-Marquardt-Levenberg algorithm. PEST is successfully applied in many fields of geophysical sciences, in particular in groundwater modelling. Its applicability for CPU intensive distributed hydrological models is focus of this study.

Alpine catchments carry specific challenges. Response times of runoff peaks after precipitation events can be extremely short owing to steep gradients. Moreover, climate zones change over short distances as temperatures decrease with higher elevations. Snow dynamics significantly influences runoff behaviour. Often, only very little information on groundwater and hydrogeology is available. All these constraints hold true for the Ammer catchment.

Both the Ammer catchment and the setup of WaSiM is described in Kunstmann et al. (2004) who investigated the impact of climate change on the hydrology of the Ammer catchment. It is repeated for reasons of completeness in the following section.

## 2 Hydrological modeling of the Ammer catchment

The Ammer catchment drains into Lake Ammersee, located around 50 km south-west of the city of Munich (Germany). The catchment covers an area of around 710 km<sup>2</sup> in southern Bavaria (Fig. 1). The landscape of the alpine and prealpine drainage basin is characterised by high spatial geological and pedological differentiation, a complex orography and corresponding specific climatological conditions. The highest elevation within the catchment is 2185 m (Kreuzspitze), and the outflow into Lake Ammersee is at 533 m above sea level.

The catchment can be divided into two landscape units: the prealpine hill country and moorland and the Swabian-Upper Bavarian foothills of the Alps. The main geological units are the limestone-alpine zone in the southern part, the flysch zone bordering in the north, the folded molasses and the unfolded molasses in the northern part of the catchment.

Soils have formed since the late glacial and post glacial (e.g. Ludwig, 2000); soil type distribution is shown in Fig. 2. An overview of recent vegetation types is given in the land use map shown in Fig. 3. It is based on the interpretation of Landsat-TM images for which a fuzzy logic algorithm was applied. Around 50% of the area is (mostly coniferous) forest, which constitutes the major vegetation unit. Digital elevation, soil and vegetation data were obtained from public data sources derived in the EU funded RAPHAEL project ("Runoff and Atmospheric Processes for Flood Hazard Forecasting and Control", Bacchi and Ranzi, 2000) in which the catchment of the Ammer River (among other European alpine catchments) was investigated.

The catchment's climate is cool-temperate and humid. Precipitation maximum is in summer. Owing to the relief all climate variables have latitude- and altitude-dependent gradients. Long-term mean temperature is 7–8°C. In the southern mountainous regions, the mean temperature drops to 4.5°C. Temperature distribution shows an altitude-dependent gradient of around 0.6°C/100 m in summer and 0.45°C/100 m in winter. Mean precipitation is around 1400 mm/year of which 67% falls in summer. In the prealpine region maximum precipitation is 140–160 mm in June; in the alpine region it is >200 mm in July. Days with snow cover (snow depth >10 cm) in the catchment are around 130 days/year. Radiation is less variable. In the prealpine part of the catchment it is around 1100 kWh/m<sup>2</sup> and rises to 1200 kWh/m<sup>2</sup> in the alpine summit regions (Ludwig, 2000).

The receiving water body of the Ammer is Lake Ammersee, which in turn drains (via the river Amper) to the River Danube. Characteristic water discharges of the Ammer catchment at its eight runoff gauges are given in Table 1. Specific discharges in the catchment are comparatively high, which is typical for alpine and prealpine environments in humid climates.



**Fig. 1.** Location of the Ammer catchment in southern Germany and its 8 sub-catchments (adapted from Kunstmann et al., 2004).

### *Model description and setup*

The distributed hydrological model WaSiM was applied (Schulla and Jasper, 2001) to simulate the hydrology of the Ammer catchment. WaSiM was chosen because of its applicability to mountainous terrains (Schulla, 1997; Jasper et al., 2002; Kunstmann and Stadler, 2005), and its assumed predictability under climate change conditions (e.g. Schulla, 1997; Kunstmann et al., 2004). WaSiM uses physically based algorithms for the majority of the process descriptions and applies conceptual approaches where physical parameters

would be difficult to obtain. It usually can be applied with standard available gridded land surface information. WaSiM uses an infiltration approach after Green and Ampt (1911), estimates saturation time after Peschke (1987), and solves the Richards equation (Richards, 1931; Phillip, 1969) for the description of the soil water fluxes in the unsaturated zone (Jasper et al., 2002). The dependence of the suction head and the hydraulic conductivity on soil moisture content is parameterized according to van Genuchten (1976). Corresponding soil type dependent parameters are based on literature values (Schulla and Jasper, 2001). Being aware

**Table 1.** Characteristic water discharges of the 8 gauges in the Ammer catchment (NQ: lowest observed discharge, MNQ: mean low water discharge, MQ: mean discharge, MHQ: mean high water discharge, HQ: highest observed discharge, MNQ: mean low water specific discharge, Mq: mean specific discharge, MHq: mean high water specific discharge).

	Gauge	Episode	Elev. [m]	NQ [m <sup>3</sup> /s]	MNQ [m <sup>3</sup> /s]	MQ [m <sup>3</sup> /s]	MHQ [m <sup>3</sup> /s]	HQ [m <sup>3</sup> /s]	MNq [l/s*km <sup>2</sup> ]	Mq [l/s*km <sup>2</sup> ]	MHq [l/s*km <sup>2</sup> ]
1	Linderhof	1983–1998	964	0.002	0.009	0.094	1.5	3.78	3.9	41.0	638
2	Oberammergau	1961–1990	831	0.4	1.2	3.6	52.5	135	10.9	31.9	462
3	Unternogg	1974–1999	849	0.2	0.3	1.7	46.4	126	7.4	38.2	1070
4	Obernach	1961–1990	652	0.01	0.1	1.0	29.2	49.2	3.2	25.1	704
5	Oberhausen	1961–1990	585	0.1	0.7	2.6	21.5	51.7	6.0	22.4	184
6	Peißenberg	1961–1990	592	1.3	2.9	8.9	114	286	10.0	30.1	387
7	Weillheim	1961–1990	550	2.6	5.7	14.7	156	338	9.4	24.5	260
8	Fischen	1961–1990	533	3.0	6.2	16.4	173	283.0	8.7	23.1	244

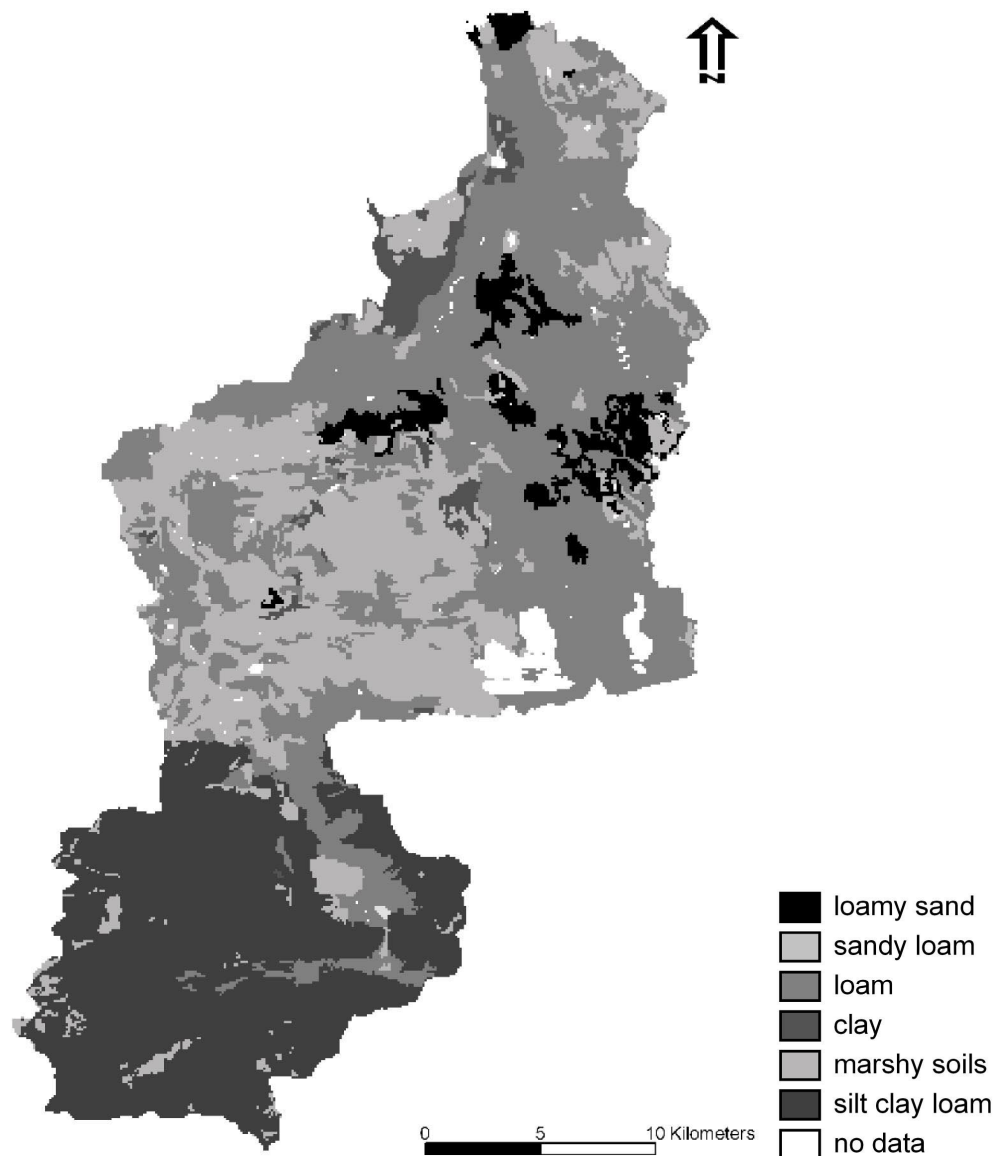
of the scale dependency of soil and vegetation parameters, the literature values are interpreted and accepted as crude approximations in this study. Unlike e.g. in Kunstmann (2005), subgrid-scale heterogeneity is not accounted for in the calculation of the water and energy balance at the land surface. Interflow is calculated in defined different soil layers, depending on suction, drainable water content ( $d$ ), hydraulic conductivity ( $k$ ), and gradient. Surface runoff is routed to the subbasin outlet using a subdivision of the basin into flow time zones. For considering retention, a single linear storage approach is applied to the surface runoff in the last flow time zone (with recession constant  $k_d$ ). Translation and retention of interflow is treated accordingly (recession constant  $k_i$ ). Potential and real evapotranspiration is calculated soil and vegetation specific using the Penman-Monteith equation (Monteith, 1975; Brutsaert, 1982). Interception is accounted for by a bucket approach. Snow accumulation and snowmelt are modelled according to Anderson (1973) and Braun (1985). Surface runoff is created for each grid cell as the sum of infiltration excess and snowmelt along the topographic gradient towards the next river. It is assumed that saturated hydraulic conductivity decreases (depending on soil texture) with depth according to a recession constant  $k_{\text{rec}}$ . Discharge routing is performed by a cinematic wave approach using different flow velocities for different water levels in the channel. After the translation of the wave a single linear storage is applied to the routed discharge considering diffusion and retention (Schulla and Jasper, 2001). WaSiM was applied with an integrated 2-D groundwater flow model, which couples dynamically to the unsaturated zone. Technical details on the coupling approach and its numerical realisation are given in Schulla and Jasper (2001). The lower boundary condition of the unsaturated zone is the depth of the groundwater layer, which is assumed to be constant for a specific time step but variable in time due to groundwater flow, recharge or capillary rise. The coupling between ground water (saturated zone) and unsaturated zone

is realised by a net vertical boundary flux. The groundwater table is located in the lowest partly saturated layer. Its exact location within the thickness of the layer is interpolated, based on the assumption of hydrostatic equilibrium between water content in the centre of the layer and the groundwater table. The uppermost (and in this study single) aquifer is assumed to be unconfined. Infiltration from rivers into groundwater and exfiltration (which is the base flow) from groundwater into rivers is calculated using the hydraulic gradient and the colmation (in- and exfiltration resistance) at the river bed. WaSiM does not solve the heat flux balance in the soil/subsurface and therefore it is not able to account for frozen soil effects.

The river courses of the Ammer catchment were derived from a digital elevation model (DEM) with  $100 \times 100 \text{ m}^2$  resolution and presumed values on mean Manning roughness and specific discharge for the channels. Flow direction correction algorithms were applied and flow orders assigned according to Strahler (1964). For each stream link a flow velocity after Manning's equation as well as surface runoff velocity is estimated.

A horizontal resolution of  $100 \times 100 \text{ m}^2$  was chosen. The unsaturated zone was parameterized with 30 layers of 0.33 m thickness. It is especially the description of the flow through the unsaturated zone and the high number of discretization layers that makes WaSiM computationally demanding: at each time step the water balance in the unsaturated zone at  $2.13 \times 10^6$  grid points is solved.

The groundwater model requires data on 1) saturated hydraulic conductivity in horizontal direction, 2) aquifer thickness, 3) specific storage coefficients and 4) leakage coefficients (or colmation resistances) to describe the exfiltration/infiltration of groundwater and river water along the river bed. Additionally, in- and outflow across the aquifer boundary must be specified, as well as constant heads for being able to solve the partial differential equation for groundwater flow. Crude information on aquifer thickness

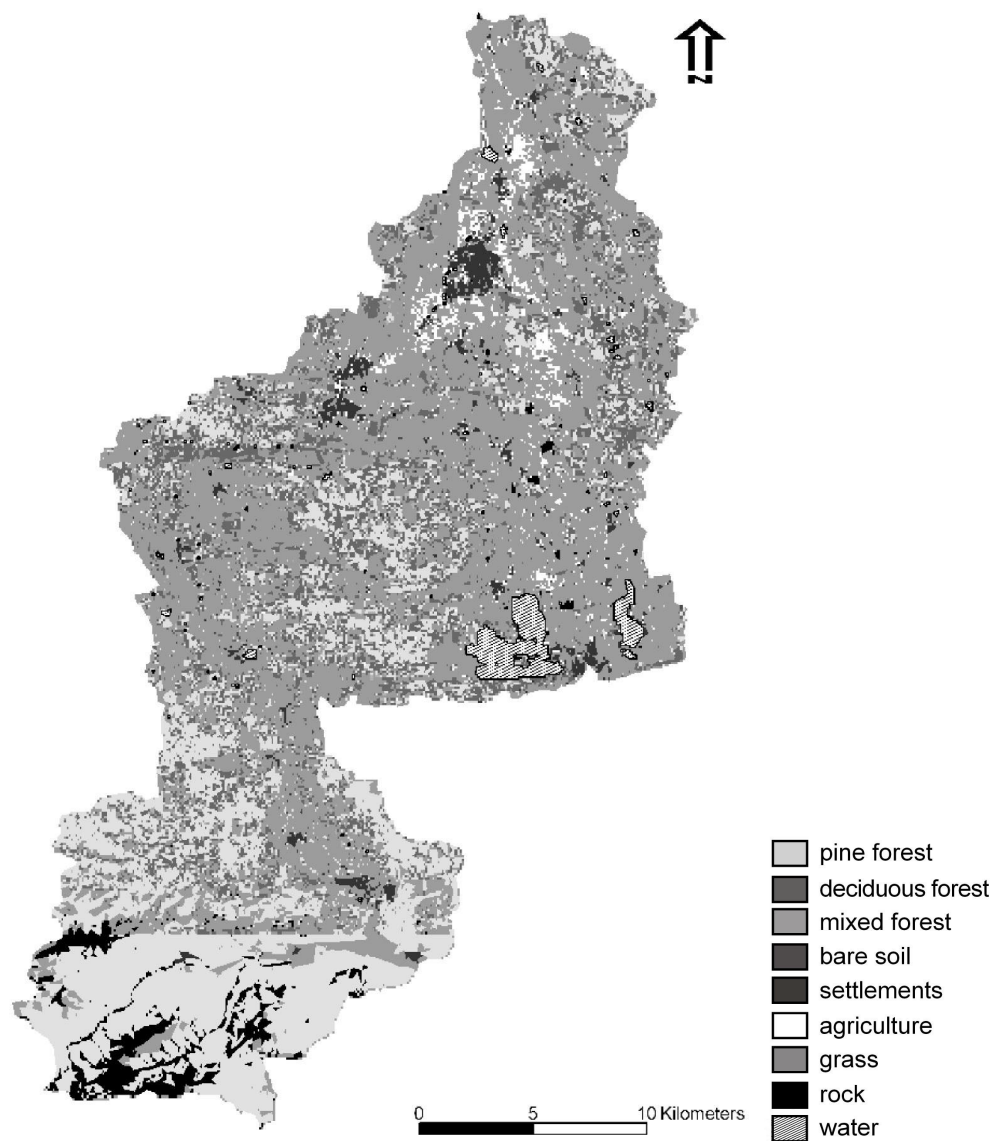


**Fig. 2.** Soil character distribution of the Ammer River catchment.

was available from a few pumping tests (Riekell, 1983). Values for storage coefficients were derived from porosity values of the soil map, as the uppermost (and single) aquifer is assumed to be unconfined. Values for colmation resistances were based on experience of similar alpine catchments such as the River Mangfall 50 km east of the Ammer (Kunstmann and Stadler, 2005) and the River Thur in Switzerland (Schulla and Jasper, 2001). Within the groundwater model, therefore, only the saturated hydraulic conductivities remained for parameter estimation. Expertise gained in the RAPHAEL project (LUDWIG, 2000) suggests the amount of groundwater outflow in the limestone-alpine zone (south-eastern part of the basin): it is estimated that around 25% of the entire runoff in sub-catchment 2 flows subterraneously

to the bordering catchment of the Loisach River owing to karstic conditions. Hence, an outflow in sub-catchment 2 of  $-9.75 \times 10^{-8}$  m/s was assumed for 120 grid cells along the south-eastern boundary. Additionally, the levels of Lake Riegsee and Lake Staffelsee were used to approximate constant heads in the area of the two lakes. This was necessary since the partial differential equations for groundwater flow require at least one constant head to determine a unique solution for piezometric heads. The choice of the level of these two lakes is justified by the fact that both lakes are in full hydraulic contact to the aquifer and show only slight fluctuations over the year.

Altogether, 8 gauges are available within the Ammer catchment (Fig. 4). Additional “fictitious” gauges were set



**Fig. 3.** Classified land use distribution of the Ammer River catchment.

to separate in- and outflow to Lake Staffelsee (gauges no. 9–11) and Lake Riegsee (gauge 12). To avoid error propagation down the stream network, calculated routed discharge is replaced by the observed value at the gauging station. This allows parameter estimation for downstream catchments independent of potential model weaknesses upstream.

Meteorological input to WaSiM was obtained by interpolation of meteorological station data (temperature, precipitation, wind velocity, humidity and sunshine duration) at 15 locations to every grid cell. The quality of interpolated precipitation is limited because none of the meteorological stations was located inside the catchment (Fig. 5). Only one station (Hohenpeissenberg) was at the edge; the remaining 14 stations were outside the catchment but within a radius of 50 km

from the centre of the catchment. Horizontal interpolation was performed by inverse distance weighting (IDW); in the case of temperature, altitude dependent regression (ADR) was applied (with changing regression parameters at every time step; derived from station data). In the case of precipitation, combined IDW and ADR interpolation was applied (with weights of 70% for IDW and 30% for ADW). Temporal resolution was 8 h.

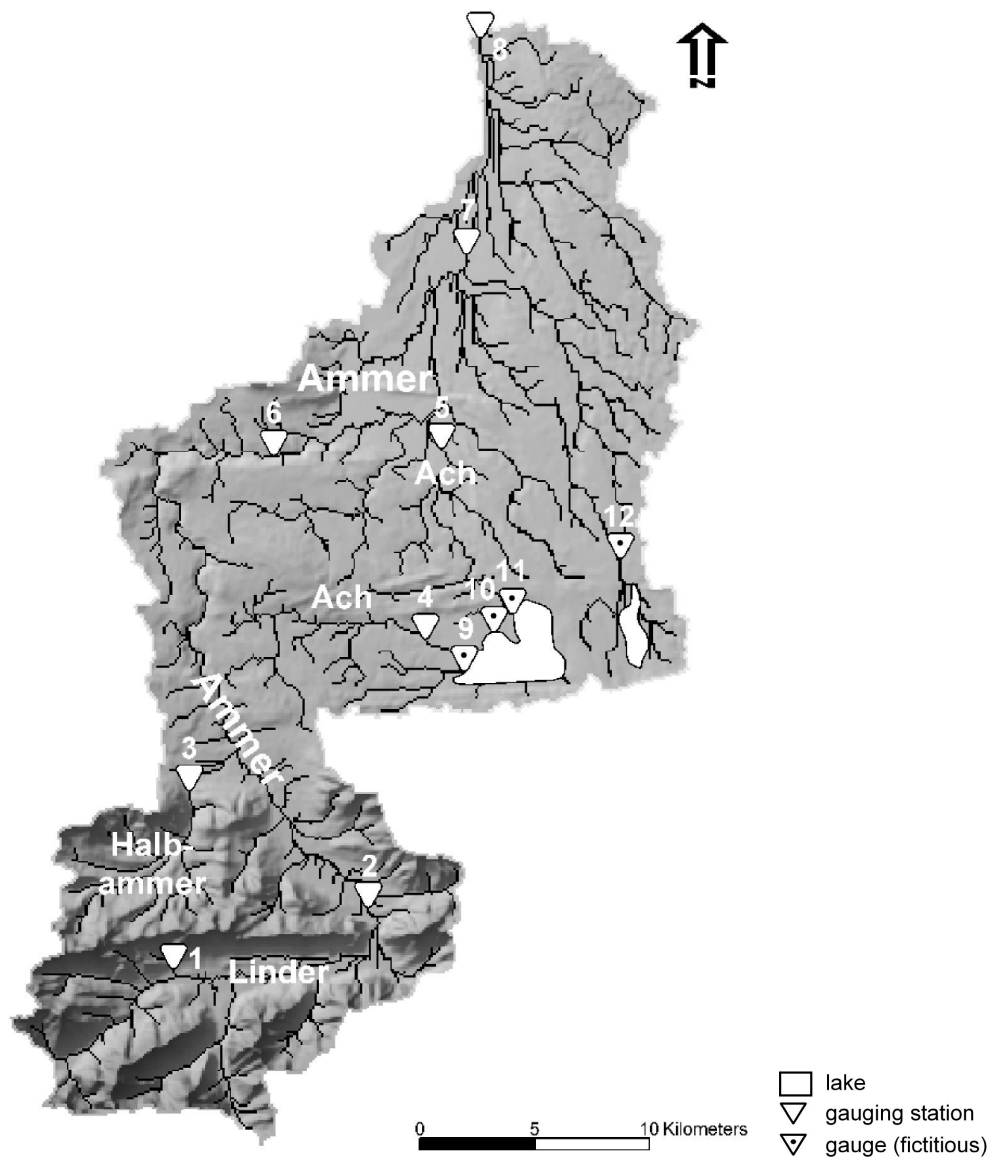


Fig. 4. Relief Map of the Ammer River catchment – incl. location of 8 gauges (adapted from Kunstmann et al., 2004).

### 3 Review: The Gauss-Marquardt-Levenberg algorithm as realized in PEST

PEST uses the Gauss-Marquardt-Levenberg algorithm for nonlinear parameter estimation. For linear models, parameter estimation can be achieved in one step. However, for non-linear problems (WaSiM falls into this category), parameter estimation can only be achieved by an iterative process. How this iterative estimation works is briefly reviewed here (following Doherty's, 2002, formulation).

The relationship between parameters and model-generated output can be represented by a function  $M$  which maps the  $n$ -dimensional parameter space into the  $m$ -dimensional “obser-

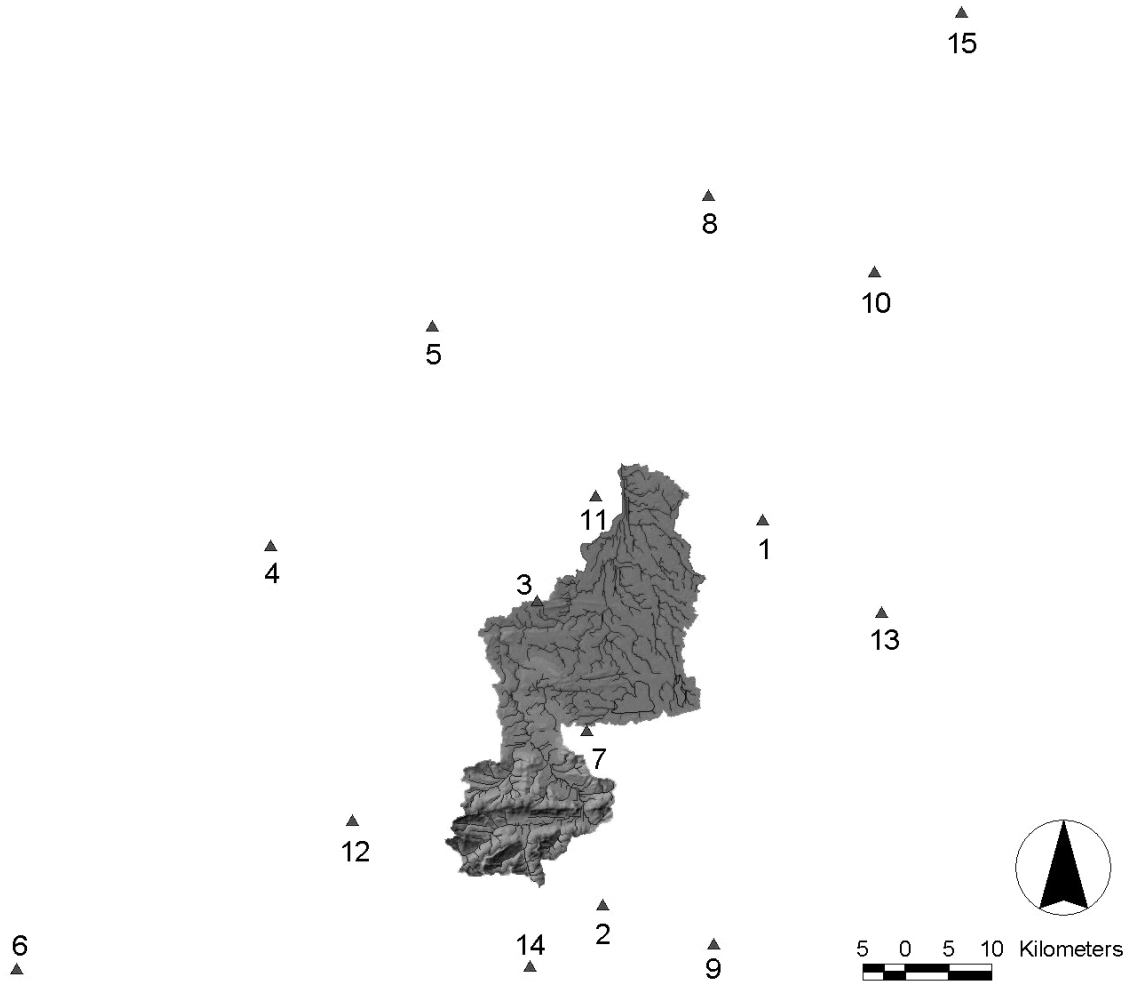
vation” space. It is required that this function is continuously differentiable with respect to all model parameters for which estimates are sought. Suppose that for the set of parameters to be estimated in the hydrological model (comprising the parameter vector  $p_0$ ) the corresponding set of model-calculated discharges is  $q_0$ , i.e.

$$q_0 = M(p_0) . \quad (1)$$

A parameter vector  $p$  that differs slightly from  $p_0$  then produces a model output  $q$  that can be approximated to (Taylor's theorem)

$$q \approx q_0 + J \cdot (p - p_0) . \quad (2)$$





**Fig. 5.** Location of the meteorological stations.

Here,  $\mathbf{J}$  indicates the Jacobian matrix of  $\mathbf{M}$ , consisting of  $m$  rows (one for each observation of model output) and  $n$  columns.  $\mathbf{J}_{ij}$  indicates the derivative of the  $i$ -th observation with respect to the  $j$ -th parameter.

Inverse hydrological modelling means that a set of model parameters is estimated for which the model generated river runoff is as close as possible to the observed runoff. In the least square sense this means that a set of parameters has to be found for which the objective function  $\chi^2$ , defined as

$$\chi^2 = (\mathbf{q}_{\text{obs}} - \mathbf{q}_0 - \mathbf{J} \cdot (\mathbf{p} - \mathbf{p}_0))^t \cdot \mathbf{W} \cdot (\mathbf{q}_{\text{obs}} - \mathbf{q}_0 - \mathbf{J} \cdot (\mathbf{p} - \mathbf{p}_0)), \quad (3)$$

is a minimum (superscript “ $t$ ” denotes the transposed matrix). Here,  $\mathbf{q}_{\text{obs}}$  represents the observed discharge values and  $\mathbf{q}_0$  the modelled discharge values.  $\mathbf{W}$  is a  $m \times m$  diagonal matrix whose entries  $w_{ii}$  are the squares of the weights attached to the  $i$ -th observation. Introducing observation weights allows higher contribution to the objective function for observations that have a higher reliability. The weights do not necessarily have to sum up to unity. Often they are chosen

inversely proportional to the standard deviation of the measurements. A new estimate for the parameter  $\mathbf{p}$  can be obtained by

$$\mathbf{p} = \mathbf{p}_0 + \mathbf{u} \quad (4)$$

with the upgrade vector

$$\mathbf{u} = (\mathbf{J}^t \cdot \mathbf{W} \cdot \mathbf{J})^{-1} \cdot \mathbf{J}^t \cdot \mathbf{W} \cdot (\mathbf{q}_{\text{obs}} - \mathbf{q}_0) \quad (5)$$

(superscript “ $-1$ ” denotes the inverse matrix). Since Eq. (2) is only approximately correct, so also is Eq. (5). Hence, the vector  $\mathbf{p}$  (defined by Eq. 4) adding the parameter upgrade vector  $\mathbf{u}$  to the current parameters values  $\mathbf{p}_0$  does not guarantee to yield the minimum of the objective function. The new set of parameters contained in  $\mathbf{p}$  must then be used as a new starting point in determining a further parameter upgrade vector and so forth.

Marquardt (1963) and Levenberg (1944) changed Eq. (5) to

$$\mathbf{u} = (\mathbf{J}^t \cdot \mathbf{W} \cdot \mathbf{J} + \alpha \mathbf{I})^{-1} \cdot \mathbf{J}^t \cdot \mathbf{W} \cdot (\mathbf{q}_{\text{obs}} - \mathbf{q}_0) \quad (6)$$

**Table 2.** Model run based upon unadjusted parameters ( $k_d=10$ ,  $k_i=100$ ,  $d=10$ ): quality of fit for the year 1997 for each sub-catchment.

	1	2	3	4	5	6	7	8
lin NS	−1.02	−3.47	−1.17	0.38	−0.43	−0.02	−0.14	0.03
log NS	0.25	−2.53	−1.84	0.52	−1.74	−2.02	−2.40	−0.49

(with  $\mathbf{I}$  denoting the  $n \times n$  identity matrix) thereby introducing a parameter  $a$ . When  $a$  is zero, Eq. (6) is equivalent to Eq. (5). When  $a$  is high, the direction of  $\mathbf{u}$  approaches that of the negative gradient vector  $\mathbf{g}$ , defined as

$$\mathbf{g}_i = \frac{\partial \chi^2}{\partial p_i} \quad (7)$$

which can be expressed as

$$\mathbf{g} = -2\mathbf{J}^t \cdot \mathbf{W} \cdot (\mathbf{q} - \mathbf{q}_0). \quad (8)$$

The advantage of this strategy is a faster convergence to the minimum of the objective function  $\chi^2$ , in particular when parameters are correlated. Details on the strategy of how PEST chooses the Marquardt parameter  $a$  can be found in (Doherty, 2002).

PEST uses a secant's approximation for approximating the Jacobian matrix  $\mathbf{J}$ . This is achieved by perturbation of the parameters to be estimated (by 1% e.g.). In fact, estimating  $n$  parameters requires  $n$  perturbed model runs and one unperturbed run, i.e.  $n+1$  model calls for every iteration within PEST. Usually, after 5 iterations no further improvements in  $\chi^2$  were reached in our case.

Based on the Jacobian matrix the sensitivity of each parameter with respect to all observations can be calculated by

$$s_i = (\mathbf{J}^t \cdot \mathbf{W} \cdot \mathbf{J})_{ii}^{1/2} / m \quad (9)$$

with  $m$ : number of observations and  $i$  indicating the number of the parameter. Relative parameter sensitivity  $rs_i$  is then defined as the product of  $s_i$  and the parameter value  $p_i$ :

$$rs_i = s_i \cdot p_i. \quad (10)$$

#### 4 Calibration strategy, results of inverse modelling and validation

Calibration of the hydrological model required adjustment of the following parameters: 1) recession constant of direct runoff  $k_d$ , 2) recession constant of interflow  $k_i$ , 3) drainage density  $d$ , and 4) saturated hydraulic conductivities of the upper most aquifer ( $k_{\text{sat}}$ ). Additionally, five snowmelt parameters had to be adjusted for the entire catchment in order to account for regional snow cover dynamics: 1) the temperature for beginning snowmelt  $T_0$ , 2) the temperature dependent melt factor  $C_1$ , 3) the wind dependent melt factor  $C_2$ , 4)

the temperature at which 50% of precipitation falls as snow  $T_{R/S}$ , and 5) the temperature transition range from snow to rain  $T_{\text{trans}}$ .

To evaluate the quality performance of the calibration, the Nash-Sutcliffe criterion NS was used (Nash and Sutcliffe, 1970):

$$\text{NS} = 1 - \frac{\sum_{i=1}^n (q_{\text{sim},i} - q_{\text{obs},i})^2}{\sum_{i=1}^n q_{\text{obs},i}^2 - \frac{1}{n} \left( \sum_{i=1}^n q_{\text{obs},i} \right)^2} \quad (11)$$

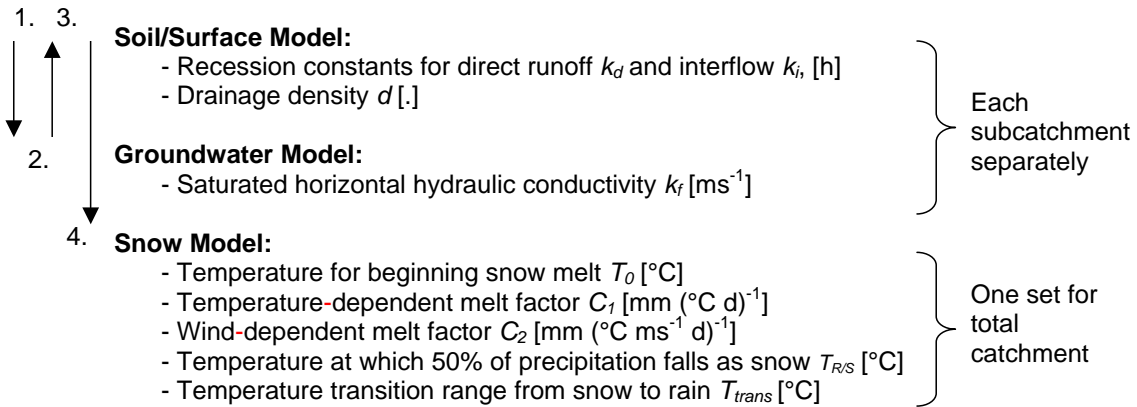
(with  $n$ : number of river runoff measurements,  $i$ : index of gauge,  $q_{\text{sim}}$ : simulated runoff,  $q_{\text{obs}}$ : observed runoff). The range of the NS values extends from 1 to  $-\infty$ . In spite of being commonly used in performance analysis, the linear NS (hereafter referred to as “lin NS”) has the disadvantage that it overstates the quality of flooding peaks against the quality of low flow periods. The performance for simulating low flow periods is better characterised by using logarithms of runoff values (indicated by “log NS” in this paper).

As explained in Sect. 5, PEST minimizes the objective function  $\chi^2$  (Eq. 3) rather than maximizing the Nash-Sutcliffe criterion. While PEST in fact minimizes the numerator of the second term in Eq. (11), the NS includes weights of this term through the denominator. It is only this weighting which PEST does not account for.

For all 8 sub-catchments, 3 surface parameters and 1 groundwater parameter had to be calibrated. Additionally, 5 parameters describing snow cover dynamics had to be estimated for the entire catchment. Altogether,  $8 \times 3 + 8 + 5 = 37$  parameters had to be calibrated. The model was given a spin-up time of 10 months (from January 1996 till October 1996) to allow the water contents in soil and rivers to be adjusted.

The calibration strategy is visualised and summarised in Fig. 6. In a first iteration step, the three surface parameters (recession constant for direct runoff  $k_d$ , recession constant for interflow  $k_i$ , drainage density  $d$ ) were calibrated for each of the 8 sub-catchments separately. For this task, the 2-dimensional groundwater model was switched off and base flow was calculated in a conceptual way (for details of the conceptual approach see Schulla and Jasper, 2001).

The quality of the model to describe observed runoff using default values and no a-priori information (i.e. fully uncalibrated state) for  $k_d$ ,  $k_i$ , and  $d$  is shown in Table 2. To get improved first guesses (“starting values”) for the order of



**Fig. 6.** Calibration strategy.

magnitude of the recession constants, a hydrograph analysis was performed by analysing the slope of the falling limb after peak flow. Hydrograph analysis is a prerequisite to force the recession constants in the proper region of the parameter space. According to the Maillet formula (e.g. Dyck and Peschke, 1995) the falling limb after peak flow can be described according to

$$Q_t = Q_0 e^{-\frac{t}{k}} \quad (12)$$

with  $Q_t$ : discharge after  $t$  days,  $Q_0$ : peak discharge at  $t=0$  days,  $t$ : time,  $k$ : recession constant. A first guess of the recession constant  $k$  can be determined by evaluation of the slope  $\ln(1/k)$  of the linear equation

$$\ln Q_t = \ln Q_0 - \ln(1/k) \cdot t \quad (13)$$

Hydrograph analysis was applied for both  $k_d$  (first part of the falling limb) and  $k_i$  (second part).

To force the Gauss-Marquardt-Levenberg algorithm into the correct physical parameter half space, boundaries for the parameter search are additionally provided. The recession constant for interflow is restricted to be larger than the recession constant for direct (“quick”) runoff. The results of the quality of fit after the first iteration are given in Table 3. It can be seen that the NS values for all sub-catchments are all positive and not lower than 0.29. However, some parameters reached the given boundary of the parameter search (such as  $k_i$  in subcatchments 3 and 4).

In iteration step 2, the 2-dimensional groundwater model is switched on. Since crude numbers for aquifer thicknesses were available, the hydraulic conductivities were calibrated. In general, a hydrogeologically motivated zonation of hydraulic conductivities would be appropriate. Since this information was not available for the Ammer catchment, however, zones for the hydraulic conductivities that were identical to the surface sub-catchments were defined. This simplification of real (but unknown) distribution of hydraulic conductivities

was motivated by the result of a sensitivity analysis (Kunstmann and Stadler, 2005) that showed that the base flow reacts rather sensitively in WaSiM to perturbation in hydraulic conductivities. Since groundwater in the aquifer crosses the “borders” of different (surface) sub-catchments, hydraulic conductivities in one sub-catchment influence base flow (and therefore the hydrograph) of other gauges as well. Calibration of the hydraulic conductivity in one sub-catchment therefore requires the quality of fit at all gauges to be considered. This was achieved by compiling the objective function  $\chi^2$  (Eq. 3) from 8 terms, each term describing the deviation of modelled and simulated runoff at a single gauge. No specific weighting of the 8 terms to the total  $\chi^2$  was performed. The quality of fit after iteration 2 is given in Table 4 and the estimated saturated hydraulic conductivities are shown in Table 5. Here, additionally log NS is provided since it is a better measure for low flow conditions (and hence base flow) than lin NS. It can be seen that application of the 2-dimensional groundwater model yielded a decline of lin NS for gauges 1, 2, 5 and 7 against the conceptual model as it was applied in the first iteration. This shows that the application of the more sophisticated sub-model does not necessarily lead to better reproduction of observed values. The estimated values for the saturated hydraulic conductivity are within physical reasonable ranges for the alpine/prealpine environment.

In iteration step 3, calibration of the surface parameters is repeated, but now applying WaSiM with the 2-dimensional groundwater model. The parameter limits were slightly adjusted. The results for estimated recession constants  $k_d$  and  $k_i$ , and the drainage density  $d$  are given in Table 6. In this iteration step, all  $3 \times 8 = 24$  parameters had to be estimated in one single PEST run. It should be noted again that each of the 8 hydrographs is used to fit 3 parameters. In this iteration step, only NS values for sub-catchments 4 and 5 are improved significantly.

In iteration step 4, the 5 snow model parameters are estimated. These parameters significantly influence snow

**Table 3.** First calibration iteration: starting values, parameter limits, fit results and quality of fit (calibration of soil/surface parameters, no 2-dimensional groundwater model applied).

Catchment		Start	Limits	Result	Nash-Sutcliffe	
					Start	Result
1	$k_d$ [h]	59.8	[5–100]	21.0	lin NS –0.15	lin NS 0.46
	$k_i$ [h]	329.7	[100–800]	304.3		
	d [.]	10	$[1 \times 10^{-8}–200]$	11.5		
2	$k_d$	121.2	[5–170]	39.6	lin NS 0.2	lin NS 0.37
	$k_i$	551.5	[170–800]	575.7		
	d	10	$[1 \times 10^{-8}–200]$	14.0		
3	$k_d$	121.6	[5–170]	19.6	lin NS 0.2	lin NS 0.45
	$k_i$	373.7	[170–800]	170.0		
	d	10	$[1 \times 10^{-8}–200]$	8.4		
4	$k_d$	138.0	[1.0–200]	1.7	lin NS 0.01	lin NS 0.51
	$k_i$	587.2	[200–800]	10.0		
	d	10	$[1 \times 10^{-8}–200]$	10.5		
5	$k_d$	167.7	[5–170]	43.8	lin NS 0.5	Lin NS 0.69
	$k_i$	424.6	[170–800]	172.8		
	d	10	$[1 \times 10^{-8}–200]$	14.6		
6	$k_d$	139.9	[5–300]	9.0	lin NS 0.24	lin NS 0.61
	$k_i$	698.9	[300–800]	38.0		
	d	10.	$[1 \times 10^{-8}–200]$	35.0		
7	$k_d$	137.4	[5–140]	17.8	lin NS 0.23	lin NS 0.55
	$k_i$	659.5	[140–800]	172.8		
	d	10	$[1 \times 10^{-8}–200]$	200.0		
8	$k_d$	165.8	[1.0–300]	18.0	lin NS –0.05	lin NS 0.29
	$k_i$	753.7	[300–800]	20.0		
	d	10.	$[1 \times 10^{-8}–200]$	8.2		

(\* Starting values were obtained by hydrograph analysis and tangent approximations for peak recessions.)

cover dynamics, and in particular the delayed transformation of winter precipitation into snowmelt dominated runoff in spring time. All five snow model parameters are restricted to comparatively narrow physical limits. The results of the calibration of the snow model parameters are given in Table 7. There is a slight improvement of the Nash-Sutcliffe values for the alpine-type sub-catchments 1 and 2, but a decline in the NS values for lower sub-catchment 7, as Table 8 shows.

It is stressed that major improvements in the quality of fit were already achieved by the first iteration step. In general, the following iteration steps provided only slight further improvements of the quality of fit. Application of the physically based 2-dimensional groundwater model yielded a decline of the NS values that could not fully be compensated in the further iteration steps. Application of the 2-dimensional groundwater model, however, has the advantage that water balance simulation comprises quantification of groundwater recharge and the interaction of groundwater and river water

**Table 4.** Second calibration iteration: quality of fit after calibration of 2-dimensional groundwater model for each sub-catchment.

	1	2	3	4	5	6	7	8
lin NS	0.24	0.33	0.47	0.34	0.04	0.53	0.51	0.15
log NS	0.46	0.19	0.34	0.62	0.39	0.30	0.47	0.16

which is important for further studies performed in the catchment.

Altogether, the Nash-Sutcliffe values and the quality of fits are reasonable but not convincing for some sub-catchments. Figures 7 and 8 show modelled and observed river runoff at gauges 4 (Oberrach) and 6 (Peissenberg) for the hydrologic year 1996/1997. Model performance in the winter months is in general lower than in the summer months. A general problem in this catchment is the fact that it contains no meteorological station. Especially in complex terrain, the pure

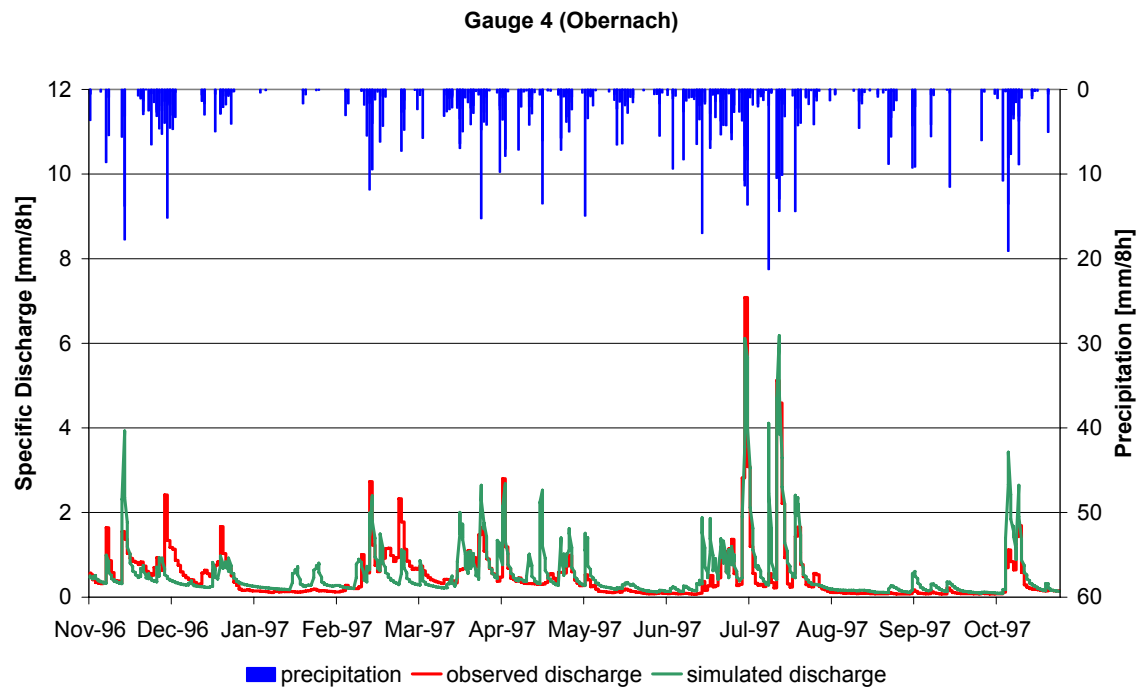


Fig. 7. Modelled vs. observed discharge at gauge 4 (Obernach, calibration period).

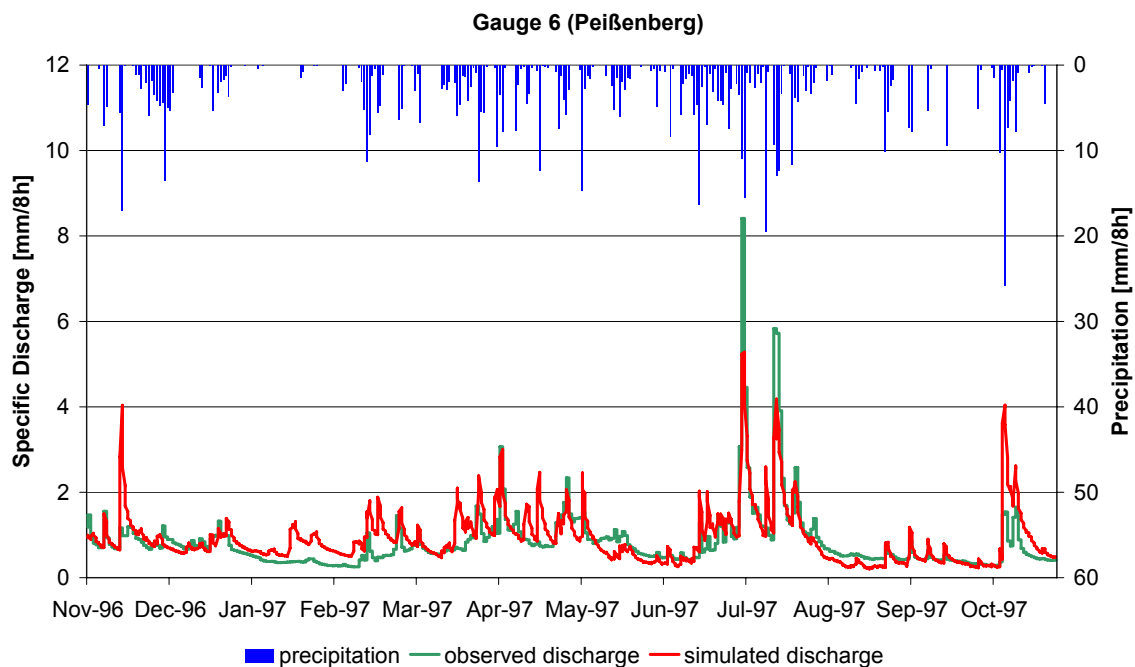


Fig. 8. Modelled vs. observed discharge at gauge 6 (Peißenberg, calibration period).

geometric interpolation (through inverse distance weighting) of precipitation across narrow valleys and ridges leads to significant errors in precipitation input to the model. The same applies for temperature interpolation, which sensitively influ-

ences snow cover dynamics. Erroneous snow cover buildup in the winter months inevitably leads to erroneous snowmelt and snowmelt-dominated runoff in spring and early summer months.

**Table 5.** Results for estimated hydraulic conductivities in groundwater model and their corresponding uncertainties (standard deviations  $\sigma$ , as described in Sect. 4.2) for each sub-catchment.

Hyd. Cond.	1	2	3	4	5	6	7	8
$k_f$ [m/s]	2.0 E-5	3.3 E-5	5.5 E-5	1.5 E-8	1.2 E-6	6.4 E-7	2.1 E-6	1.7 E-5
$\pm\sigma$	9.6 E-7	6.2 E-7	6.2 E-7	3.2 E-7	2.4 E-7	2.8 E-7	1.3 E-7	4.3 E-7

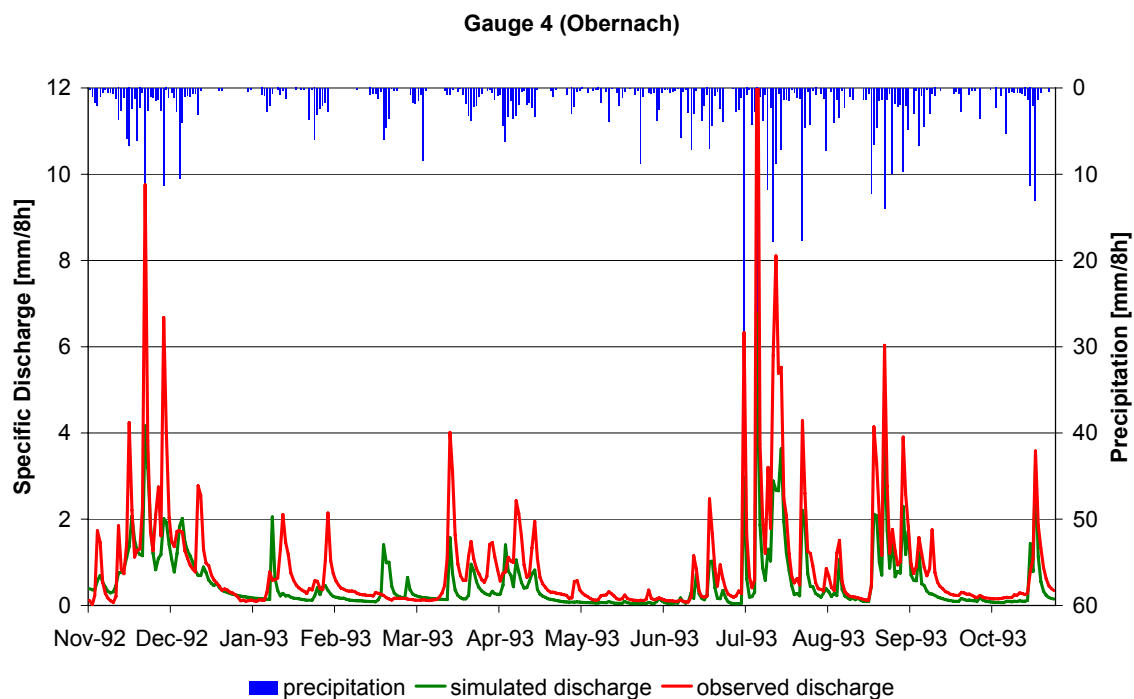
**Table 6.** Third calibration iteration: starting values, parameter limits, fit results and quality of fit (calibration of soil/surface parameters, 2-dimensional groundwater model applied). The uncertainty ranges (standard deviations  $\sigma$ ) according to covariance analysis are also provided (details see Sect. 4.2).

Catchment		Start	Limits	Results	$\pm\sigma$	Nash-Sutcliffe	
1	$k_d$ [h]	21.0	[5–100]	24.5	1.0	lin NS 0.24	log NS 0.46
	$k_i$ [h]	304.3	[100–800]	420.9	47.1		
	$d$ [.]	11.5	$[1 \times 10^{-8} - 200]$	10.5	0.1		
2	$k_d$	39.6	[5–100]	55.7	3.0	lin NS 0.35	log NS 0.19
	$k_i$	575.7	[100–800]	676.3	87.8		
	$d$	14.1	$[1 \times 10^{-8} - 200]$	14.5	0.1		
3	$k_d$	19.7	[0.5–10]	23.5	0.9	lin NS 0.48	log NS 0.32
	$k_i$	170.0	[100–800]	223.9	27.5		
	$d$	8.4	$[1 \times 10^{-8} - 200]$	8.4	0.1		
4	$k_d$	1.7	[1.0–15]	10.0	1.0	lin NS 0.55	log NS 0.65
	$k_i$	10.0	[10–800]	19.6	2.7		
	$d$	10.6	$[1 \times 10^{-8} - 200]$	11.4	0.1		
5	$k_d$	43.9	[5–100]	100.0	13.8	lin NS 0.48	log NS 0.47
	$k_i$	172.8	[100–800]	509.7	108.0		
	$d$	14.7	$[1 \times 10^{-8} - 200]$	14.8	0.15		
6	$k_d$	9.1	[1–15]	15.0	1.3	lin NS 0.62	log NS 0.49
	$k_i$	38.0	[15–800]	95.8	7.8		
	$d$	35.0	$[1 \times 10^{-8} - 200]$	35.6	0.3		
7	$k_d$	17.8	[5–100]	13.9	4.6	lin NS 0.52	log NS 0.47
	$k_i$	172.8	[100–800]	164.5	13.4		
	$d$	200.0	$[1 \times 10^{-8} - 200]$	194.8	1.7		
8	$k_d$	18.0	[1.0–20]	20.0	1.7	lin NS 0.32	log NS 0.25
	$k_i$	20.0	[20–800]	190.3	111.3		
	$d$	8.2	$[1 \times 10^{-8} - 200]$	8.0	0.1		

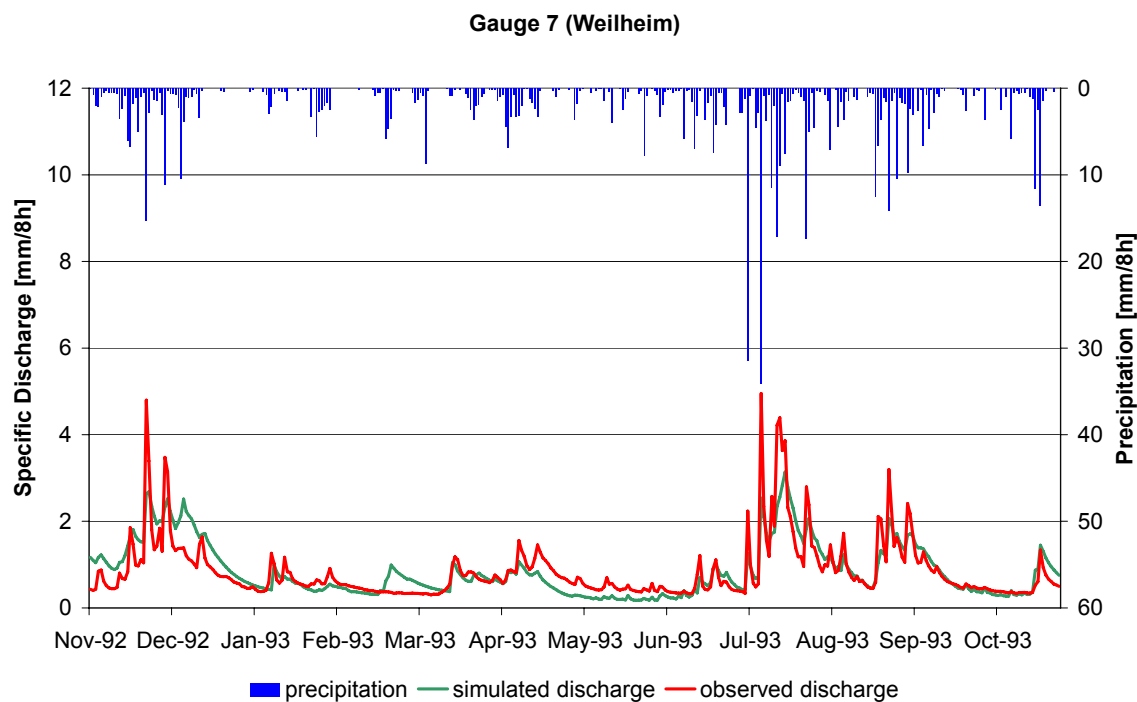
#### 4.1 Validation

The quality of the simulations for the validation period (hydrologic year 1992/1993) can be seen from Table 9. Compared to the calibration period, the lin NS values are lower for the Alpine sub-catchments 1, 2, 3 and 4, but higher for the sub-catchments 5, 7, and 8 of the downstream part of the Ammer catchment. The log NS values are in general lower in the validation period, except for sub-catchment 7. Due to the differences in changes of lin NS and log NS values from the calibration phase to the validation phase, it can be concluded that the derived parameters are better designed

for capturing high-flow situations than low-flow conditions. This corresponds to the fact that high discharge values receive more weight in the objective function (Eq. 3) than low flow discharge values. Figures 9 and 10 show modelled and observed river runoff exemplary at gauges 4 (Oberrach) and 7 (Weilheim) for the validation period. Likewise in the calibration period, the model performance in the winter months and in the spring time is in general lower than in the summer months, due model deficiencies in describing snow accumulation and snow melt dynamics and corresponding impact on surface runoff.



**Fig. 9.** Modelled vs. observed discharge at gauge 4 (Obernach, validation period).



**Fig. 10.** Modelled vs. observed discharge at gauge 7 (Weilheim, validation period).

**Table 7.** Results for estimated snow-model parameters, including uncertainty ranges (standard deviations  $\sigma$ ).

Snowmelt parameter	Start	Limits	Result	$\pm\sigma$
$T_0$ [°C]	0.6	[−0.5–0.8]	0.26	0.02
$C_1$ [mm/°C*d]	0.3	[0.001–1.75]	0.61	0.02
$C_2$ [mm/°C*m/s*d]	0.5	[0.001–1.5]	0.12	0.01
$T_{R/s}$ [°C]	0.6	[0.001–1.0]	0.29	0.03
$T_{trans}$ [°C]	1.5	[0.001–2.0]	0.79	0.09

**Table 8.** Fourth calibration iteration: final quality of fit (calibration of snowmelt parameters included) for each sub-catchment.

	1	2	3	4	5	6	7	8
lin NS	0.28	0.42	0.51	0.57	0.49	0.64	0.45	0.34
log NS	0.43	0.22	0.32	0.64	0.42	0.52	0.37	0.26

#### 4.2 Derivation of confidence bounds of estimated parameters

Confidence bounds for all parameters estimated can be obtained by covariance analysis. The covariance matrix of the estimated parameters can be approximated by

$$[\mathbf{Cov}] \equiv [\boldsymbol{\alpha}]^{-1} \quad (14)$$

with

$$\alpha_{kl} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial p_k \partial p_l} \approx \sum_{i=1}^N \left[ \frac{\partial q(t_i, \mathbf{p})}{\partial p_k} \frac{\partial q(t_i, \mathbf{p})}{\partial p_l} \right] \quad (15)$$

(Press et al., 1992).

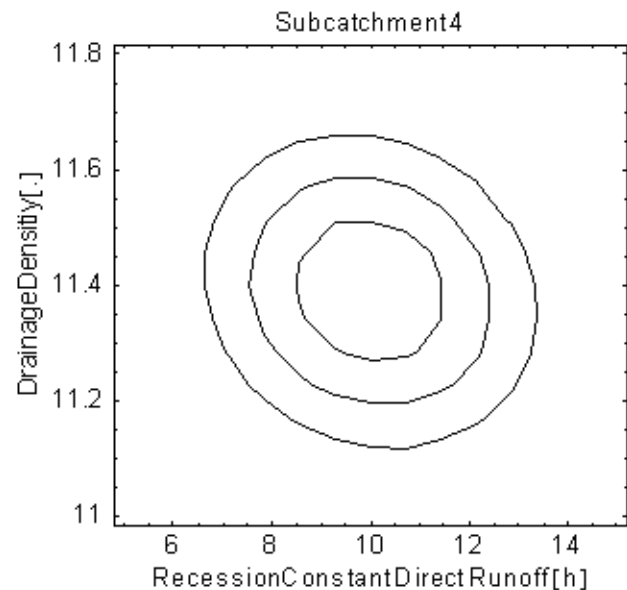
The square roots of the diagonals of the parameter covariance matrix yield the standard deviations of the estimated parameters quantifying the uncertainty of the estimated values.

Tables 5, 6 and 7 include values for the estimated uncertainty range (standard deviation  $\sigma$ ) of all estimated parameters as derived from Eqs. (14) and (15). It can be seen that the uncertainty bounds are comparatively narrow, except in case of recession constants for interflow which show comparatively high uncertainty ranges. This reflects a small sensitivity of  $\chi^2$  with respect to the recession constants for interflow which in turn is more difficult to estimate.

The standard deviations describe the uncertainty range only in case of negligible correlation between the estimated parameters. The correlation of the  $n$  estimated parameters makes the solution space an  $n$ -dimensional hyper-ellipsoid. If this hyper-ellipsoid is projected into 2-dimensional subspaces, confidence ellipses are obtained. The confidence ellipses are defined by (e.g. Press et al., 1992) (superscript

**Table 9.** Period of validation: quality of fit for the hydrologic year 1992/1993 for each sub-catchment.

	1	2	3	4	5	6	7	8
lin NS	−0.33	0.03	0.23	0.45	0.70	0.44	0.66	0.60
log NS	−0.45	−0.35	0.05	0.28	0.06	−1.06	0.47	0.04

**Fig. 11.** Confidence ellipses for estimated parameters  $k_d$  (recession constant for direct runoff) and  $d$  (drainage density) for sub-catchment 4. The isolines indicate the confidence regions 68.3% (innermost ellipse), 95.4% and 99.7% (outermost ellipse).

“−1” indicates the inverse matrix, superscript “ $T$ ” the transposed matrix)

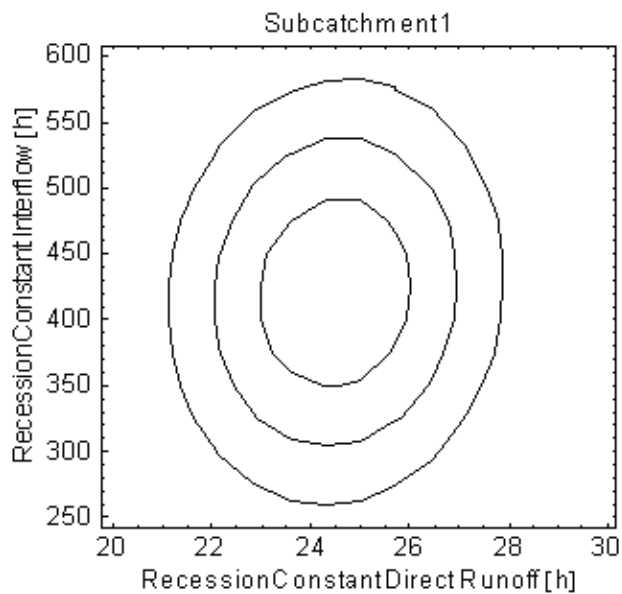
$$\Delta \Xi^2(\mathbf{p}) = (\mathbf{p} - \mathbf{p}_{\text{bestfit}}) \cdot [\mathbf{Cov}]^{-1} \cdot (\mathbf{p} - \mathbf{p}_{\text{bestfit}})^T \quad (16)$$

(Press et al., 1992) with “Cov”:  $2 \times 2$ -dimensional projection of the  $n \times n$  parameter covariance matrix,  $\Delta \Xi^2$  a tabulated function depending on confidence level (e.g. Press et al., 1992),  $\mathbf{p}$  the 2-dimensional vector of parameters of interest, and  $\hat{\mathbf{p}}$  the 2-dimensional vector of parameters of the mean values of the parameters of interest. Applying Eq. (16) to the derived Hesse matrix  $\boldsymbol{\alpha}$  and the estimated parameters  $\mathbf{p}_{\text{bestfit}}$ , confidence ellipses for all combination of parameters can be delineated. If two parameters are correlated, the principle axes of the ellipsoid are rotated with respect to the parameter axes (for more details on the analysis of confidence regions see Carrera and Neumann, 1986). Figure 11 shows the confidence regions for drainage density and recession constant of direct runoff for sub-catchment 4. In this specific case, slight correlation between the model parameters drainage density and recession constant is revealed. The isolines indicate the confidence regions for the estimated parameters. In our



**Table 10.** Relative parameter sensitivities for bestfit parameter set according to Eq. (10).

	1	2	3	4	5	6	7	8
$s(k_d)$ [.]	1.3 E-03	9.8 E-04	1.3 E-03	5.4 E-04	3.8 E-04	6.1 E-04	1.6 E-04	6.4 E-04
$s(k_i)$ [.]	4.8 E-04	4.1 E-04	4.3 E-04	4.0 E-04	2.5 E-04	6.6 E-04	6.5 E-04	9.2 E-05
$s(d)$ [.]	7.8 E-03	7.9 E-03	7.4 E-03	7.7 E-03	7.3 E-03	7.8 E-03	7.6 E-03	7.4 E-03
$s(k_f)$ [.]	1.8 E-01	2.2 E-02	4.7 E-02	1.8 E-05	2.2 E-05	9.5 E-04	2.9 E-03	1.7 E-02
$s(T_0)$ [.]				6.3 E-04				
$s(C_1)$ [.]				1.9 E-03				
$s(C_2)$ [.]				8.8 E-04				
$s(T_{R/S})$ [.]				4.5 E-04				
$s(T_{trans})$ [.]				1.8 E-04				

**Fig. 12.** Confidence ellipses for estimated parameters  $k_d$  (recession constant for direct runoff) and  $k_i$  (recession constant for interflow) for sub-catchment 1. The isolines indicate the confidence regions 68.3% (innermost ellipse), 95.4% and 99.7% (outermost ellipse).

examples, the isolines indicate the probabilities of 68.3% (innermost ellipse, equivalent to 1 standard deviation  $\sigma$ ), 95.4% (equivalent to  $2\sigma$ ) and 99.7% (outermost ellipse, equivalent to  $3\sigma$ ) to find the real parameter within the corresponding confidence region. Figure 12 shows confidence regions for the estimated recession constants for direct runoff and interflow in sub-catchment 1. As already revealed in the pure correlation-neglecting covariance analysis, the uncertainty for recession constants is comparatively large, while uncertainty of recession constants for direct runoff and drainage density is smaller. A detailed analysis of parameter correlation, however, showed that correlation is in most cases negligible. It is therefore concluded that it is appropriate to provide the standard deviation alone, as is done in Tables 5, 6 and 7.

Application of this methodology allows the detailed quantification of uncertainty of all estimated parameters and of correlation among estimated parameters. It would exceed the scope of this paper to show all combinations of confidence ellipses. Therefore, we stress the potential of the applied parameter estimation method (as it is realized in the tool PEST): the results of the confidence bounds could be used as input to stochastic hydrological modelling, e.g. by means of Monte Carlo simulations. In this way, the uncertainty propagation of input parameter uncertainty into the corresponding range of water balance variables (surface runoff, evapotranspiration, etc.) can be quantified. Estimation of input parameter uncertainty is an essential prerequisite for this task.

In addition to confidence ellipses, parameter sensitivities are of particular interest as they quantify the influence on model performance. Moreover, the higher the sensitivity of a model parameter, the more reliable is the estimated value of the parameter. The relative sensitivities of model parameters in the case of the best fit parameters are shown in Table 10. The relative parameter sensitivity is in general one order of magnitude higher for the drainage densities  $d$  than for the recession constants  $k_d$  and  $k_i$ . This is in congruency with Fig. 11, which shows much smaller confidence bounds for  $d$  than for  $k_d$ . The relative sensitivities for the snow parameters are in the same order of magnitude as the soil/surface parameters. The comparatively high relative sensitivity of the hydraulic conductivities for specific sub-catchments stresses the importance of fitting aquifer variables to improve modelled discharge.

## 5 Summary and conclusions

We combined the distributed hydrological model WaSiM (Schulla and Jasper, 2001) and the parameter estimation tool PEST (Doherty, 2002) which applies the Gauss-Marquardt-Levenberg algorithm for the nonlinear estimation of model parameters. Using PEST WaSiM was calibrated for the alpine catchment of the Ammer River in southern Germany. For each of the 8 sub-catchments 3 surface parameters and

one groundwater parameter had to be calibrated, as well as 5 snow model parameters for the entire catchment. Altogether, 37 parameters had to be calibrated. Owing to the complexity and the high spatial resolution of the hydrological model the estimation of all parameters was an extensive numerical experiment. Parameters were estimated by a four-step iteration approach, in which lumped surface parameters were estimated separately from aquifer parameters and snow-model parameters. Constraints on lower and upper boundaries were added to the parameter estimation process to force the algorithm towards the correct parameter half-space. Analysis of confidence bounds showed small uncertainties for most of the estimated parameters except for the recession constants for interflow. Correlation between the estimated parameters was negligible.

The quality of modelled discharges, both in the calibration period and the validation period, was reasonable but not always satisfactory. It proved difficult to reproduce runoff periods that are influenced by snow cover dynamics. Discrepancies between modelled and observed runoff were also due to the small number of meteorological stations and corresponding interpolation artefacts in the orographically complex terrain. Judging the comparatively overall low lin NS and log NS values, it must be considered that a CPU intensive physically and grid based distributed hydrological model was applied rather than a conceptual, lumped model. Parameter estimation methods that are applied successfully for conceptual models and that require large number of model calls cannot be applied here (e.g. several thousands of model runs in case of simulated annealing or genetic algorithms for one parameter). Considering additionally the large number of parameters to be adjusted, it is obvious that the calibration of this type of physically based hydrological model does not necessarily reach the quality of quick lumped conceptual models. The complexity of the hydrological model (accounting for high resolution distributed information on all aspects of the surface/subsurface water balance) is required for understanding the physical water balance in the catchment. It is assumed to be in particular suited for applications which require predictability capabilities, like e.g. requested in climate change impact studies (as e.g. presented in Kunstmann et al., 2004, for the same Ammer catchment). However, the complexity of the model has the trade-off that NS values are smaller than in case of conceptual models. Lower quality of NS values may have to be accepted under this circumstance.

It was particularly shown that it was not possible to achieve a higher model performance when applying a complex, dynamically coupled 2-dimensional numerical groundwater flow model instead of a conceptual simple model approach. Higher complexity did not provide higher model performance. Nevertheless, as discussed above, the dynamically coupled surface water – groundwater modelling approach is of importance for specific research questions like the impact of climate change or land use change on terrestrial water balance components and particularly groundwater resources.

Due to this fact it is currently investigated whether parameter sets that satisfactorily describe river discharge also reasonably reproduce observed heads and vice versa. This is achieved by joint use of river discharge and groundwater information in the automatic parameter estimation.

While not neglecting the above mentioned constraints, this work showed that the applied Gauss-Marquardt-Levenberg algorithm is suited to calibrate a CPU intensive distributed, mostly physically based hydrological model such as WaSiM. It is concluded that the proposed parameter estimation methodology can advantageously be coupled to similar complex and CPU-time intensive hydrological models. Moreover, since the method provides covariances and uncertainty estimates on the estimated parameters it is a prerequisite for stochastic hydrological modelling such as Monte Carlo simulations, for which uncertainties (e.g. through standard deviations as derived in this study) on input parameter are required.

*Acknowledgements.* This work was partially performed within the framework of the Centre of Numerical Environmental Simulation and funded by the Bavarian High Tech Initiative. Their financial support is gratefully acknowledged. We greatly appreciated the help of J. Schulla and K. Jasper in applying WaSiM. Special thanks also go to M. Hannweber of the Wasserwirtschaftsamt Weilheim, who provided runoff data and local expertise.

Edited by: E. Zehe

## References

- Aarts, E. H. L. and Korst, J.: Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing, John Wiley and Sons, Chichester, 1989.
- Abdulla, F. A., Lettenmaier, D., and Liang, X.: Estimation of the ARNO model baseflow parameters using daily streamflow data, *J. Hydrol.*, 222, 37–54, 1999.
- Anderson, E.: National Weather River Forecast System – Snow Accumulation and Ablation Model, NOAA, Tech. Mem., NWS-Hydro-17, U.S. Department of Commerce, USA, 1973.
- Bacchi, B. and Ranzi, R. (Eds.): RAPHAEL – Runoff and Atmospheric Processes for Flood Hazard Forecasting and Control, Final report to the EC, Directorate General XII, Programme Environment and Climate 1994–1998, 2000.
- Bastidas, L., Gupta, H., Sorooshian, S., Shuttleworth, W., Yang, Z.: Sensitivity Analysis of a Land Surface Scheme using Multi-Criteria Methods, *J. Geophys. Res.*, 104(D16), 19 481–19 490, 1999.
- Braun, L. N.: Simulation of snowmelt-runoff in lowland and lower alpine regions of Switzerland, *Zürcher Geographische Schriften*, ETH Zürich, 21, 166 pp, 1985.
- Brutsaert, W.: Evaporation into the Atmosphere, Kluwer Academic Publishers, Dordrecht, 1982.
- Carrera, J. and Neuman, S.: Estimation of Aquifer Parameters under Transient and Steady State Conditions: 1. Maximum Likelihood Method Incorporating Prior Information, *Water Resour. Res.*, 22, 199–210, 1986.

- Dickinson, R., Henderson-Sellers, A., and Kennedy, P.: Biosphere Atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model, NCAR Technical Note, NCAR, 72 pp, 1993.
- Doherty, J.: PEST – Model-Independent Parameter Estimation, Watermark Numerical Computing, Australia, 2002.
- Doherty, J. and Johnston, J. M.: Methodologies for calibration and predictive analysis of a watershed model, *J. Amer. Water Resour. Assoc.*, 39(2), 251–265, 2003.
- Doherty, J.: PEST, PUB-IAHS Workshop “Uncertainty Analysis in Environmental Modeling”, 6–8 July 2004, Lugano ([http://www.es.lancs.ac.uk/hfdg/uncertainty\\_workshop/uncert.methods.htm](http://www.es.lancs.ac.uk/hfdg/uncertainty_workshop/uncert.methods.htm)), 2004.
- Dyck, S. and Pescke, G.: Grundlagen der Hydrologie, Verlag für Bauwesen, Berlin, 1995.
- Duan, Q., Gupta, V. K., and Sorooshian, S.: Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28(4), 1015–1031, 1992.
- Duan, Q., Gupta, V. K., and Sorooshian, S.: Optimal use of the SCE-UA global optimization method for calibrating watershed models, *J. Hydrol.*, 158, 265–284, 1994.
- Green, W. H. and Ampt, G. A.: Studies on Soil Physics: I. The flow of air and water through soils, *J. Agric. Sci.*, 4, 1–24, 1911.
- Hundecha, Y. and Bárdossy, Á.: Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, *J. Hydrol.*, 292, 281–295, 2004.
- Jasper, K., Gurtz, J., and Lang, H.: Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model, *J. Hydrol.*, 267, 40–52, 2002.
- Kunstmann, H. and Stadler, C.: High Resolution Distributed Atmospheric-Hydrological Modelling for Alpine Catchments, *J. Hydrol.*, 314, 105–124, 2005.
- Kunstmann, H.: Upscaling of land surface parameters through direct moment propagation, *Adv. Geosci.*, 5, 127–131, 2005.
- Kunstmann, H., Schneider, K., Forkel, R., and Knoche, R.: Impact analysis of climate change for an Alpine catchment using high resolution dynamic downscaling of ECHAM 4 time slices, *Hydrol. Earth Syst. Sci.*, 8, 1030–1044, 2004.
- Levenberg, K.: A method for the solution of certain non-linear problems in least squares, *Quart. J. Appl. Mathematics*, 2, 164–168, 1944.
- Ludwig, R.: Die flächenverteilte Modellierung von Wasserhaushalt und Abflußbildung im Einzugsgebiet der Ammer, *Münchner Geographische Abhandlungen*, Band 32, München, 2000.
- Marquardt, D.: An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Indust. Appl. Mathematics*, 11, 431–441, 1963.
- Monteith, J. L.: *Vegetation and the Atmosphere*, Vol. 1: Principles, Academic Press, London, 1975.
- Nash, J. E. and Sutcliffe, J. V.: River Flow Forecasting through Conceptual Models. Part I – a Discussion of Principles, *J. Hydrol.*, 10(3), 398–407, 1970.
- Peschke, G.: Soil Moisture and Runoff Components from a Physically Founded Approach, *Acta hydrophysica*, 31(3/4), 191–205, 1987.
- Phillip, J. R.: The theory of infiltration, in: *Advances in Hydrosciences*, edited by: Chow, V. T., Academic Press, New York, 216–296, (1969).
- Press, W., Teukolsky, S., Vetterling, W., and Flannery, B.: *Numerical Recipes in C*, Cambridge University Press, Cambridge, 1992.
- Richards, L. A.: Capillary Conduction of liquids through porous medium, *Physics*, 1, 318–333, 1931.
- Riekkel, T.: Hydrogeologische und hydrochemische Untersuchungen im Flußgebiet der Ammer, Diss., TU München, München, 1983.
- Schulla, J.: Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen, Diss., ETH 12018, Verlag Geographisches Institut ETH Zürich, 187 p., 1997.
- Schulla, J. and Jasper, K.: Model Description WASIM-ETH (Water Balance Simulation Model ETH), ETH-Zurich, Zurich, 2001.
- Strahler, A.: Quantitative geomorphology of drainage basins and channel networks, in: *Handbook of Applied Hydrology*, edited by: Chow, V. T., Maidment, D. R., and Mays, L. W., McGraw-Hill, New York, 1964.
- Van Genuchten, M. T.: A Closed Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, *Amer. J. Soil Sci.*, 44(5), 892–898, 1976.
- Vrugt, J., Gupta, H., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, 39(8), 1201, 2003a.
- Vrugt, J., Gupta, H., Bastidas, L., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resour. Res.*, 39(8), 1214, 2003b.
- Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-Objective Global Optimization for Hydrologic Models, *J. Hydrol.*, 204, 83–97, 1997.